

Supporting Statement For OMB Review and Approval of

Agency for Toxic Substances and Disease Registry (ATSDR)
A Prospective Birth Cohort Study Involving Environmental Uranium Exposure
in the Navajo Nation

SECTION B. Statistical Methods

Date: 1 February 2013

Program Official:
Angela Ragin-Wilson, PhD
Agency for Toxic Substances and Disease Registry
Division of Toxicology and Human Health Sciences
Program Official
Phone: 770-488-3807
Fax Number: 770-488-7187
Email: atr0@cdc.gov

Point of Contact:
Candis M. Hunter, MSPH
Agency for Toxic Substances and Disease Registry
Division of Toxicology and Human Health Sciences
Phone: 770-488-1347
Email: chunter@cdc.gov

Table of Contents

B. STATISTICAL METHODS

- B.1. Respondent Universe and Sampling Methods
- B.2. Procedures for the Collection of Information
- B.3. Methods to Maximize Response Rates and Deal with Nonresponse
- B.4. Test of Procedures or Methods to be Undertaken
- B.5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

B.1. Respondent Universe and Sampling Methods

Respondent Universe & Background:

The study participants will be pregnant women who receive prenatal care or are willing to receive prenatal care from the following Navajo Area Indian Health Service (NAIHS) /NAIHS Public Law 93-638 (P.L. 93-638) contracted facilities: Northern Navajo Medical Center, Chinle Comprehensive Health Care Facility, Gallup Indian Medical Center, Tuba City Regional Health-Care Corporation, or Tséhootsooí Medical Center. NAIHS is one of 12 area offices of the Indian Health Service, a branch of the U.S. Public Health Service of the Department of Health and Human Services. The NAIHS is responsible for the delivery of health services through a combination of in-patient, out-patient and community health programs centered on six hospitals, seven health centers, and 15 health stations. Delivery of service is accomplished through eight service units. Navajo Nation administers care in some of the service units (i.e., Tuba City, Tséhootsooí , Kayenta) through P.L. 93-638 contracts administered by Navajo Nation Division of Health. NAIHS clinicians document approximately 3,000 Navajo live births per year in the greater NAIHS system, distributed among the participating hospitals as indicated Table B.1 below.

Table B.1: Navajo Area OB/GYN Statistics (2009) Self-Reported at Area and National Meetings (Source NAIHS clinical staff)

Service Unit	Chinle (NAIHS)	Tséhootsooí (PL-638)	Gallup (NAIHS)	Northern Navajo(NAIHS)	Tuba City (PL-638)	Total
Number of Births	541	459	664	763	519	2946

The NAIHS is responsible for the delivery of health services through a combination of in-patient, out-patient and community health programs centered on six hospitals, seven health centers, and 15 health stations. The five service units in the table comprise the hospitals where the majority of deliveries will occur for Navajo Nation, although high-risk deliveries may be referred to larger regional hospitals such as the University of New Mexico or the University of Arizona. The five primary facilities encompass three IHS Service Unit Hospitals, and two that were formerly IHS facilities that have utilized PL-638 to compact services out of IHS in recent years. The five facilities chosen also represent four satellite facilities, and handle the majority of deliveries. While prenatal care can come from any of the listed facilities and records will still be accessible to the team, the delivery hospitals are those that both catch the majority of exposure areas as well as deliver the greatest number of babies each year.

The five service units (excluding Kayenta) represent the majority of the deliveries on the reservation. These service units typically have about 3,000 births a year compared to the 4,500 births in the Navajo Nation (67%).

The service units will be the targets for recruitment in this study, with recruitment criteria requesting the participants to seek care and plan delivery at one of these hospitals. These service units also correspond well to the areas of highest anticipated risk. Kayenta also serves an exposed population, and will be included in the recruitment, although deliveries from that area will occur in a neighboring hospital such as Tuba City, Shiprock, or Chinle as Kayenta does not provide that services for labor and delivery.

The most recently published Navajo birth data showed 13,634 total births in 1999-2001, which corresponds to a rate of 20.8 births per 1000 population, or approximately 4,500 births per year (USDHHS, 2008). However, in discussions with NAIHS clinicians, the figure of approximately 3,000 Navajo live births per year in the NAIHS system has been consistently cited, and may reflect what has been a declining birth rate in recent years. (For example, the number of New Mexico resident Navajo births dropped 15.1% between 1990 and 1999.) A Navajo Community Health Assessment reported 3,003 live births in NAIHS in FY 2003 and noted that approximately 15-20% of live births on the Navajo Nation were born to adolescent mothers (<18 years of age) (Benally, 2005). We used the conservative estimate of 3,000 births per year in our sample size considerations and power calculations. In our discussions to solicit input from community members and clinicians, the necessity of including these minor mothers in the cohort has repeatedly been identified. Inclusion of mothers below the age of 18 in the research will require consent from a parent for participation of the minor mom (or dad), assent from the minor for their participation, and the mom's consent for the participation of her unborn child (for which purpose she is seen as emancipated).

Case Definition: The intent is to study the possible association between exposure to uranium and either adverse birth outcomes in the case of parental exposure, or developmental delays associated either with parental exposure, in utero exposure, or exposure of the child during development. We anticipate all of these circumstances will be identified in the cohort, and will recruit women who have lived on Navajo lands for at least 5 years and fathers who have conceived a child, and are willing to have that child followed at birth, 2 months, 6 months, 9 months and up to 1 year of age. Miscarriage and mortality are important health points of interest (Please see chart in Part A for all health points of interest). However, due to the infant mortality rate and miscarriages that may occur in this population, we may not be able to get further information at 2 months up to a year of age.

Women will be identified through extensive outreach both in communities and at chapter meetings, as well as through clinical settings. We will enroll women who have lived on Navajo Nation or within adjacent communities for at least 5 years; are willing to deliver at one of the five participating clinical service units; and are willing to consent for follow-up of their child during a minimum of the first year of life. The study staff consists of the UNM research field staff, cohort clinical liaisons (CCLs), and Community Environmental Health Research Specialists (CHERs) will actively recruit participants. The study staff will explain the benefits of

being in the study to eligible women they encounter during the course of various outreach activities. The study staff will also administer screening eligibility questionnaires to interested individuals.

Fathers will be enrolled to the extent identified in order to assess environmental exposure, previous reproductive, and medical histories. Exposures of all three members of the family unit will be determined through surveys, environmental sampling and existing data assessment, and verified through biomonitoring. Please refer to “Exposure Assessment under Items of Information to be Collected” in Part A.

We are encouraging participation from the full extent of the Navajo Nation. The five designated hospitals handle the vast majority of deliveries and reach beyond any specific chapters. Uranium mining features vary greatly across the Navajo Nation. There are some parts of the reservation with many uranium mining and milling structures such as in the Shiprock, NM area. While other areas may not have as many uranium mine structures. The size of mines varies from a few feet to many acres. The water concentrations for uranium in local wells vary from non-detects to >300 ug/L (10 times the MCL). For arsenic, the range is from ND to >400 ug/L (40 times the MCL). For soil uranium concentration, it ranges from 1ppm (1mg/kg dry wt) to >500 ppm. For the related waste contaminant Ra-226, soil concentrations ranging from <1pCu/g - >800 pCu/g have been documented. Therefore, recruitment across the whole of Navajo should allow for substantial variability in the range of exposures and support our plans for developing a Bayesian Model Averaging Regression model that will be supported by a full range of exposures.

Inclusion Criteria: Native American mothers from age 14 to 45 with confirmation of pregnancy will be recruited into the study. Inclusion criteria will be a willingness to deliver at one of the five health-care facilities participating in the study, to have resided in the Study Area for at least 5 years, and to be willing to have the child followed for up to 1 year. Following the mother’s consent, the initial page of the intake survey will ask questions about the father, his knowledge of her participation, his willingness to participate, and a name for contact, if appropriate. The mother will continue to be included in the study should the father not be identified at that time. Efforts will be made to enroll eligible women as early in their pregnancy as possible (ideally, in the first trimester), but we will enroll women who volunteer for the study who are in their third trimester. The study will allow mothers who have multiple children during the study period to enroll multiple children.

The program acknowledges the importance of language consistency across the entire package with regard to the impact of prenatal care on eligibility. Prenatal care utilization is an important predictor of birth outcomes and increased utilization is an applied public health goal of this study. Participants will not be required to receive all of their prenatal care at one of the designated IHS/PL-648 study facilities in order to participate in the study, as this may be viewed as culturally insensitive. It is well documented that some Navajo mothers may not utilize the IHS facilities for care throughout their entire pregnancy. Some Navajo mothers may also choose to

visit traditional healers and/or medicine men during their pregnancy. These mothers will still be eligible to participate in the study. The enrollment survey includes questions (31-32) to ascertain whether the participant is receiving prenatal care at an IHS clinic or from a traditional healer/medicine man. Therefore, we should be able to capture prenatal care information and use it to define a metric for the Adequacy of Prenatal Care, which will be a variable in our structural equation and multivariate models.

Although willingness to deliver at one of the study hospitals is an inclusion criterion, participants are not required to receive prenatal care in order to be eligible for the following reasons:

1) IHS beneficiaries routinely “shop” for care across the facilities that are available to them. It is extremely likely that participants in this study will NOT receive all of their prenatal care at the facility where they plan to deliver. By having the broad network of participating clinical sites involved, and through regular review of participants, we are trying to design a tracking process to not lose those individuals. The reality of the situation means that we are aiming to have at least an initial visit for prenatal care at a participating facility, and to obtain participant commitment for delivery. At delivery admissions, the clinic staff will be requesting the prenatal care charts if the care was outside of the system, and those agreements are in place and supported by the participants agreement to grant access to medical records. The research team’s experience of working with clinical staff and with our research staff in the field give us confidence that we can get participants into one of the study clinics for the initial consenting and blood/urine collection, and back at delivery. In the interim, we have no control and will rely on chart reviews.

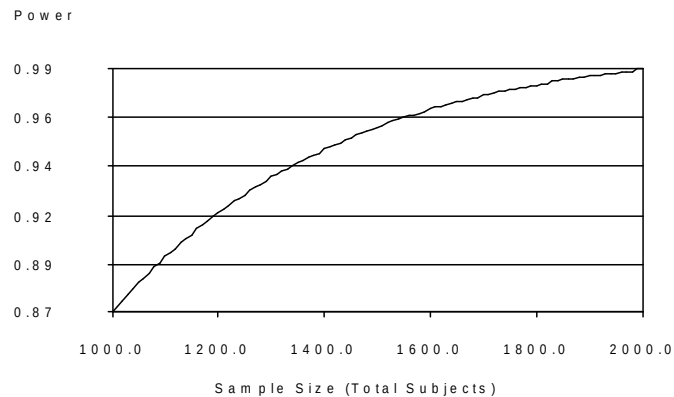
2) Collection and processing of samples is more difficult at delivery. We want participants to deliver at one of the designated study hospitals, but we realize that we may not be able to attain all samples at this point. We have designed this study to meet the realities of the situation on the ground. We accept the fact that some variables will result in less than 1500 complete records. Our final analyses will be tiered to look at the subset of complete records, as well as falling back on the alternative exposure variables as necessary to capture the largest sample size available for each endpoint. As stated, multiple analytic approaches will be used to achieve consistency in the analyses.

Fathers will be included in the study with consent regardless of age or residence. All babies born to qualifying mothers will also be included in the study with the mother’s consent. If either parent is under 18, parental consent for their participation will also be necessary, as well as their assent as a minor. Mothers who turn 18 during the course of the study will be re-consented before any further data is collected from them. Because fathers will provide data only at enrollment, they will be re-consented when they reach 18. This will occur only in cases where the father is the primary caregiver.

Exclusion Criteria: Interested participants who do not have a clinical confirmation of pregnancy will be excluded from the study. Participants who qualify for the study, but are not willing to deliver at a participating study facility, will be excluded. Women who are non-Native Americans will also be excluded from participating in the study. The inclusion criteria allows for Navajo women and other Native American women living on the Navajo Reservation to be enrolled. The legal Navajo blood quantum to be considered Navajo is 25% or at least one grandparent has to be Navajo. Other Native American women who live on the Navajo Reservation may include Hopi or Pueblo women. Inclusion criteria allow these women to be enrolled, however they will be excluded if they do not live on the Navajo Reservation. Women who are not willing to have their child followed for up to one year will be excluded. In addition, women not currently residing on the Navajo Nation or in immediately adjacent communities will be excluded. For example, women living on Hopi land and Pueblo land will be excluded from the study. All participants must have lived within the boundaries of Navajo Nation at least five years. All exclusion criteria are consistent with the intent of the study the possible association between reproductive and developmental endpoints and uranium exposures on the Navajo Nation. Living outside of the Navajo Nation, or not being pregnant, will not meet the intent. Previous cohort recruitment on the Navajo Nation resulted in fewer than 5 out of 1,304 being non-Native American, a number not sufficient to break out in an analysis. Finally, NAIHS and associated PL-638 health-care providers will be supporting the study through medical record reviews and provision of standard clinical care. Therefore, participants should be eligible to receive health care in those facilities.

Power Calculations: UNM estimates that a reasonable sample size for this study is 500 live-births per year (17% of the anticipated 3,000 Navajo births/year), for a period of 3 years giving a total anticipated sample size of 1,500 mother-infant pairs. Please refer to Chart 1 for statistical power at different sample sizes.

Chart 1: Power as a Function of Sample Size



The projections of statistical power and sample size were based on several factors and assumptions. First, since there are multiple outcomes and exposures, with different underlying relationships, these calculations are only as an example of the study power that we would expect to achieve given different numbers of study participants (total number of subjects). We assume a comparison based on using the highest quintile (20%) of exposure, as the “exposed” group, and assume the exposure will double the risk of an adverse event, for example low birth weight defined as <10th percentile of weight for-length (probability of adverse event=10%). Given these assumptions, at a total sample size of 1,500, we have 95% power to detect an effect. If we enroll even 1,000 mother-infant pairs, we would still have 87% power to detect an effect of this magnitude. However, given that the effects may be more subtle, we acknowledge that these calculations likely represent a minimum sample size. Note that it is rare in environmental epidemiology to see a doubling in effect, yet strong causal associations between been adverse birth outcomes and lifestyle and health outcomes (smoking, alcohol use during pregnancy). Furthermore, we note that the power calculations do not take into account our ability to detect differences among subgroups. We acknowledge that the sample size may not be adequately powered to factor in all of the variables identified in Figures A and B in Supporting Statement Part A, particularly when individuals are removed due to variable-specific missing data. However, the study should provide important insights regarding the relative roles of different types of factors that potentially affect birth outcomes and development in this population. As discussed in Part A, CDC is committed to working with its partners to set realistic expectations with its study participants as well as clearly articulating the limitations of the study design and, accordingly, of its conclusions when disseminating the findings to its stakeholders (including other Federal agencies, Congress, and the Navajo Nation).

Target Enrollment Estimation:

The target enrollment to achieve the desired number of mother-child pairs at 1 year of age was estimated to be an average of 550 pregnant women per year, assuming a 10% pregnancy loss. Due to several cultural and logistical factors, we cannot determine with certainty the attrition rate. The Navajo Birth Cohort Study is the first prospective epidemiologic study of pregnancy and neonatal outcomes in a uranium-exposed population; therefore, there is not a representative population that could be used to estimate attrition rates. However, the CHERs will be in frequent contact with participants to schedule follow-up meetings, appointments and conduct home interviews. Included in Section B.3 is detailed information regarding maximizing response rates and dealing with non-response.

All indications are that we will receive significant funding for this study in FY13 and likely at funding levels similar to the first 3 years. With this knowledge, we are comfortable in stating that we will have the resources to recruit and follow 1650 mother/child pairs. .If out year funding beyond FY13 is provided, the research team will consider enrolling more participants as this would yield greater power. However, at this time it is neither reasonable to try and predict if

funding will become available many fiscal years from now (and the level of that funding) nor how that funding will be applied (i.e. continuing follow up of the existing cohort for a longer period of time or dedicating resources to increase cohort size).

We are confident that we can enroll the target number of participants because we have a positive history of working with the NAIHS facilities and private hospitals on Navajo for many years, with direct experience working with Gallup and Shiprock-Northern Navajo Medical Center (NNMC) service units in recruiting participants for the medical monitoring as well as the DiNEH biological sampling programs. In addition, we anticipate our outreach plan will result in a similar willingness to participate in target service units, including Gallup and Shiprock-NNMC, each of which manages about 700-750 OB cases per year and has a significant exposed population. The Navajo population is stable; meaning those enrolled during early pregnancy will likely not be lost to the study by migration out of the study area. By partnering with the NAIHS and NNDOH, Public Health Nurses (PHNs) and Community Health Representatives (CHRs) who are already in place, we will be able to capitalize on data collected as part of routine prenatal care which will require record abstraction but not additional study visits. This will help to increase our efficiency. If funding is available, we will enroll and follow mothers and their infants for as long as is necessary

B.2. Procedures for the Collection of Information

Recruitment of Study Participants:

Active recruiting will be conducted by research field staff, CCLs and the CHERs. Awareness of the study and recruitment of study participants will be conducted through a targeted culturally appropriate approach:

1. Study pamphlets available tribal chapter meetings, chapter houses, NAIHS/PL-638 hospitals & clinics, and WIC offices (Recruitment materials and brochures available in Attachment 6 and Attachment 10)
2. Vignettes on You Tube/ Facebook
3. Informational Sessions

Study staff will introduce the project in tribal chapter meetings in partnership with the regional staff and CHERs. At that time, all risks and benefits of participation in the cohort will be introduced. New CHERs will partner with existing staff for this process as an element of training, and CHRs will also participate. We will work with NAIHS midwives and obstetricians in participating service units to distribute information on the study, and inform participants in the NAIHS medical monitoring program intakes about the work as well. Multiple venues for enrollment and consent will be available: chapter houses and WIC offices where project staff

will have outreach information, at home visits attended by CHERSs and at hospital and clinics staffed by Navajo Nation Division of Health and NAIHS personnel.

Malcolm Benally, a Navajo filmmaker and Navajo language expert, is developing a series of video vignettes that will introduce the project to the community. The vignettes are 1.5 to 3 minutes, and will consist of interviews with community members and Navajo officials, document the history of uranium mining on the Navajo Nation, and describe the purpose of the study. The vignettes also include interviews with community members and Navajo leaders including Navajo Nation Vice President Rex Lee Jim discussing the importance of the study to the Nation. The vignettes will be in English and in Navajo, and be available through Facebook and YouTube. Loops of these video materials will also be shown in hospital waiting rooms.

Recruitment will also occur through information sessions at all public events, health fairs, and other fairs on Navajo Nation, and training of staff in the Growing in Beauty, WIC, and other social support programs to provide information materials through their programs has also been implemented. When women express interest in the study at the hospital, the CCLs will provide them a miniature copy of the consent home to discuss with their family and to view the materials on the study before committing to participate.

Sixty to seventy percent of Navajos stated that they would use traditional medicine for diagnosis of illness (Benally, 2005). However, clinicians, community advisors, and traditional medicine men with whom we have spoken have confirmed that nearly all Navajo mothers will also seek prenatal care in NAIHS (including PL-638 facilities). Therefore, prenatal clinics will provide the best opportunity to identify and recruit potential participants. Through outreach of our diverse team at chapter meetings, chapter houses, and WIC offices, we will hope to reach others who might not initially have sought prenatal care at the clinics. Identification of interested women will be conveyed to the field staff by NAIHS referrals, by the CHRs, or through self-responses to information obtained at meetings, from printed materials, or radio announcements. Participants will be asked initially to complete a screening document to verify pregnancy and other eligibility criteria. This screening for eligibility can be administered by a broad range of participating agency staff, including CHRs, Women Infants and Children (WIC) personnel, Health Educators (HEs), and public health nurses (PHNs). If eligibility is confirmed, participants will be referred to Navajo Birth Cohort Study dedicated staff for consent to participate in the cohort study and for review of their medical records. Separate consents will be administered for the mother, the father, and in cases of minor parents, a parent of minor in addition to minor assents. Consents will be administered by dedicated research staff working either for NAIHS, Navajo Nation Division of Health (NNDOH), or University of New Mexico (UNM) Team.

B.3. Methods to Maximize Response Rates and Deal with Nonresponse

The research team includes a substantial number of community residents from the study area communities, as well as the collaborative. Although 60-70% of Navajos stated they would use traditional medicine for diagnosis of illness (Benally, 2005), clinicians, community advisors, and traditional medicine men with whom we have spoken have all confirmed that nearly all will also seek prenatal care in NAIHS (including P.L. 638 facilities) through the midwife programs in each service unit. Therefore, these prenatal clinics will provide the best opportunity to identify and recruit potential participants. Outreach will also be through meetings at chapter houses or public events, through word-of-mouth contact with Navajo Nation Division of Health Community Health Representatives or other Navajo Division staff members working in the communities, as well as the many community members working in IHS facilities or as part of the research team.

Within most service units, the prenatal clinics are staffed by midwives. The midwife programs are not centralized, but rather maintained independently within each service unit, and in some cases staffed in conjunction with family practice physicians. Obstetricians are on call and provide consultation or a higher level of care when complications are identified. The number of pregnant women who present at NAIHS facilities during the first trimester, however, appears to be low. Although one report suggests more than 80% of women has their initial visit in the 1st trimester (Benally, 2005), more recent indications are that the 1st trimester number is a lower 61% (USDHHS, 2008). Because those not seeking care in the first trimester may be at greater risk, this factor is a potential bias that must be addressed in the recruitment strategy by outreach to increase enrollment and prenatal care or controlled in the statistical analysis. The number of prenatal visits varies dramatically, but more than one-half of pregnant women on Navajo are seen for NAIHS prenatal care more than nine times during their pregnancy. PHNs, interpreter/drivers in the PHN program, and CHRs are often requested to contact those who miss appointments in efforts to get them to the clinics. Generally, these interventions are done in cases where specific concerns have been noted, or after repeated “no-shows”. CHERSs will be working within the CHR administrative structure and therefore will be well integrated into that system. For women who do not appear for scheduled appointments, the CHERS can conduct home interviews to ensure regular data collection at critical study time intervals as well as coordinate with CHRs to get participants to clinic appointments. Another role of the outreach will be to increase awareness of the need to meet appointments in both the mind of the mom as well as her extended family and friend’s network.

Study staff will introduce the project in chapters in partnership with the regional staff and CHERSs. This will be the initial step of the train-the-trainer model employed through all phases of community work in the project. At that time, all risks and benefits of participation in the cohort will be introduced. New CHERSs will partner with existing staff for this process as an element of training, and CHRs will also participate. We will work with NAIHS midwives and obstetricians in participating Service Units to distribute information on the study, and inform participants in the NAIHS medical monitoring program intakes about the work as well. Multiple

venues for enrollment and consent will be available: by our field staff, by CHERSs in home visits, or at hospital and clinic sites by NAIHS project staff. Identification of interested women will be conveyed to the field staff by NAIHS referrals, by the CHRs, or through self-responses to information obtained at meetings, from printed materials, or radio announcements. Participants will be asked initially to complete a screening document to verify pregnancy and other eligibility criteria (see Attachment 3h). This screening for eligibility can be administered by a broad range of participating agency staff.

4. Data analysis approach

Data will be analyzed using standard statistical software packages (R, SAS) by the UNM analytic team in collaboration with CDC's Geographic Research, Analysis and Services Program (GRASP) spatial analysts and Dr. Ettinger. The DiNEH Project modeling team has worked to develop analytically sound statistical models to incorporate multidimensional contributors to the observed outcomes. These models utilized geospatial methods and assessed appropriateness of Bayesian and determinist approaches, ensuring that all models are validated and utilize existing information to update assumptions and refine models as the datasets have grown in complexity and depth. Dr. Curtis Miller, a mathematician and statistician, and Glenn Stark, a PiBBS Fellow (Program for Interdisciplinary Biological and Biomedical Sciences), who has led this effort as part of his Ph.D. work have been included in this study team and will work closely with the GRASP team to establish appropriate geospatial models of the same rigor in analysis of these data. Dr. Ettinger has extensive training and experience in modeling health effects of environmental exposures and will be consulted on statistical analysis with the UNM analytic team.

Independent & Dependent Variables. Please see Figure A and Figure B in Part A for a graphic on independent variables (uranium exposure measured through biomonitoring, surveys, and home assessments) and dependent variables (reproductive and developmental outcomes). Figure A summarizes the inputs and their place in the model for evaluating *reproductive outcomes*. Figure B summarizes inputs, modifiers, and outcomes used to assess *developmental outcomes*. Both diagrams illustrate how diverse data sources will be used to measure inputs, modifiers (confounders, covariates), and their relative contribution to the documented outcome.

Descriptive Statistics. Univariate and bivariate summary statistics and distributional plots will be examined for all variables to identify potential outliers and influential variables. Initial analyses will describe the distribution of the exposures and outcomes in the study sample. Outcomes that are not normally distributed will be appropriately transformed before analyzing their association with uranium exposure variables and other predictors. Bivariate correlations between variables will be calculated as appropriate. Nonparametric methods (e.g., Lowess) which make no assumptions about the functional form of the relationship between continuous variables, will be used to determine the interrelationships between the exposure variables (Schwartz, 1993), and

will be used to test specific hypotheses about effects of predictor variables on outcomes, specifically in cases where transformations may not be effective to obtain normal distributions. Any potential outliers will be identified and verified for accuracy by study staff against the original data. When appropriate, regression diagnostics (such as use of residual plots and leverage plots to determine heteroscedasticity and curvilinearity), extreme studentized deviate (to detect outliers), Cook's D (to determine leverage of a particular covariate measurement on the model as a whole), and correlation matrices of the estimated coefficients (to check for collinearity) will be examined. Regression analyses with these identified points set aside will also be considered as a sensitivity analysis to determine if a few values have substantial influence on the analysis.

Insuring data integrity. From our experience in modeling other data sets in Navajo Nation, we also anticipate the need to assess analytical problems presented by outliers, missing data, and failures of participants to understand questions. A variety of internal check routines will be set up to routinely check the data set for these potential problems and to ensure the most appropriate solutions are applied prior to analysis. These may include simple subroutines to ensure the sum of years in prior and current addresses equals the age of the participant or mapping to ensure identified water sources are within driving radius of a residence. These may range, based on the situation, to imputation of data, re-contact of participants for clarification, dropping of specific participants from phases of the analyses where no statistically and scientifically justifiable resolution is available.

Covariate Selection. Potential confounding variables that may be associated with the exposures and outcomes will be identified from prior studies based on biological plausibility and will be included as covariates in multivariate models, as appropriate. We will use linear, logistic, or Cox proportional hazards regression, as appropriate, to examine bivariate associations between the key covariates and the outcomes as well as the key covariates and the exposure variables to assess potential for confounding. We will use the 10% “change-in-estimate” approach to identify statistically significant confounding variables. Some of the covariates will have strong correlations with each other and overly collinear variables will not be included jointly, to avoid producing unstable model results. If this is the case, the correlations will be noted in the reported analyses. Please see Attachment 13: Survey Question Sources and Relevance, Attachment 14: Biomonitoring Table & Analyte Justification, as well as Figures A & B in Part A for more information about covariates.

Multivariate Modeling. Multivariate modeling will be based primarily on biological rather than statistical considerations. That is, for each outcome, we will identify a priori those covariates that should be entered into the model as confounders, based on biological considerations and the current state of the literature. These variables will be kept in all regression models regardless of significance. We will also identify other variables that might be confounders, but whose connection to the outcome is not established based on biological knowledge or prior

studies. Those variables will be denoted as potential confounders. Sensitivity analyses will examine how the associations change as these other covariates are introduced into the model. Missing data and loss to follow up are inevitable issues to address in any epidemiological study, particularly when longitudinal follow up is involved. We will first consider linear or logistic regression models, as appropriate for our outcome of interest. We will also use generalized additive models (GAMs) and/or penalized regression splines to assess the appropriateness of the linear effect assumptions, or whether the effects follow some other pattern, such as a threshold. Using GAMs, the response variables will be modeled as the sum of smooth functions of the continuous independent variables and linear functions of the categorical independent variables (Hastie and Tibshirani, 1990). The effects can be summarized graphically or by computing a predicted score for different levels of exposure. This model assumes that the effects of the independent variables are additive.

Sensitivity analyses will examine how the associations change as these other covariates are introduced into the model. Missing data and loss to follow up are inevitable issues to address in any epidemiological study, particularly when longitudinal follow up is involved.

We will first consider linear or logistic regression models, as appropriate for our outcome of interest. We will also use generalized additive models (GAMs) and/or penalized regression splines to assess the appropriateness of the linear effect assumptions, or whether the effects follow some other pattern, such as a threshold. Using GAMs, the response variables will be modeled as the sum of smooth functions of the continuous independent variables and linear functions of the categorical independent variables (Hastie and Tibshirani, 1990). The effects can be summarized graphically or by computing a predicted score for different levels of exposure. This model assumes that the effects of the independent variables are additive.

Generalized mixed effects models will be used for repeated outcome measures. We will measure, to the extent possible, characteristics of subjects who are approached to participate in the study, but decline, in order to assess the potential for non-responder bias. Standard statistical methods for missing data will be used to adjust for missing data (covariates and outcomes). Structural equation modeling (SEM) may be used to evaluate the complicated interrelationships between the covariates. SEMs are used for partitioning the variance in a set of interrelated multivariate outcomes into that which is due to direct, indirect, and covariate (exogenous) effects (Succop et al., 1998). This method allows for a variety of possible outcomes based on the hypothesized relationships of the variables in the model and recognizes that the equation errors in each of the outcome variables may be correlated. In addition, SEMs allow for the inclusion of latent (unmeasured) variables in the causal model pathway. Structural equation analysis also allows for correction of measurement error in exposure variables, incorporation of multiple outcomes and incomplete cases. Using SEMs we can also incorporate hierarchical (multilevel) modeling techniques using the structural equation analysis framework. Because we will be measuring

environmental exposures at the community level, within families across several generations, as well as the individual level, it will be important to account for the resulting structure of the data.

In addition to SEMs, there are a variety of different approaches available for hierarchical modeling, including hierarchical spatial models, mixed effects, and generalized estimating equations (GEE) approaches (Heagerty et al., 1994). Bayesian as well as classical statistical methodologies can be applied to these modeling approaches, as has been done with the models utilized to understand the complexity of contributing variables in the DiNEH study. A variety of methods of analyzing complex data are applicable, and it is anticipated that the final project design will draw upon in-house expertise working closely with GRASP. Final analyses will use spatial modeling to distinguish between potential random effects associated with geospatial differences in the data not related to exposures – i.e., outcomes that may vary among specific population clusters geographically without relation to waste or exposure.

B.5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

The following individuals are members of the study team and were consulted on statistical aspects of the design.

Name & Title	Contact information	Expertise	Study contribution
Robert Annett, Ph.D. Professor of Pediatrics Dept. of Pediatrics UNM School of Medicine	RAnnett@salud.unm.edu (505) 272-6854	Consultant on National Children’s Study; pediatric neuropsychology	Survey instrument selection, analysis
Adrienne Ettinger, ScD, MPH Yale University, Assistant Professor, Department of Epidemiology & Public Health	Adrienne.ettinger@yale.edu (203) 785-6232	Designing, implementing epidemiologic birth cohort studies; participation in research design	Co-Investigator; Data analysis
Curtis Miller, Ph.D. UNM-CEHP Statistician and modeler	CMille02@salud.unm.edu (505) 272-4048	Statistics and modeling	Designed data collection; Data analysis
Candis Hunter, MSPH Agency for Toxic Substances and Disease Registry, Division of Health Studies	chunter@cdc.gov (770)488-1347	Environmental epidemiology	Power calculations and analysis
Janet Cragan, MD, MPH National Center on Birth Defects and Developmental Disabilities	jcragan@cdc.gov 404-498-3807	Child and Maternal Health	Power calculations and analysis

Myra Tucker, MPH National Ctr For Chronic Disease Prevention & Health Promotion Division Of Reproductive Health	mtucker@cdc.gov 770-488-6267	Reproductive and Maternal Health	Tribal health statistics, reproductive health outcomes
---	---------------------------------	-------------------------------------	--